

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:52:14 ; Search time 170.72 seconds
(Without alignments)
17.898 Million cell updates/sec

Title: US-09-331-631a-1_COPY_29_73

Perfect score: 252

Sequence: 1 SEFDROEYECCKRQCQMLET.....RCVSQCDKRFEDIDMSKYD 45

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR66:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	77.5	30.8	509	2	alpha-globulin typ
2	77.5	30.8	588	1	alpha-globulin B p
3	72	28.6	605	1	alpha-globulin typ
4	70.5	28.0	566	2	vicilin precursor
5	67.5	26.8	47	2	arginine/glutamate
6	67	26.6	810	2	protein PV100 (imp
7	67	26.6	1170	1	thrombospondin 1 p
8	67	26.6	1170	2	thrombospondin 1 p
9	64	25.4	554	2	grpase activating
10	63	25.0	524	2	62K sucrose-binding
11	59.5	23.6	3078	2	variant-specific s
12	58	23.0	242	2	hypotheical prote
13	58	23.0	1421	2	hypotheical prote
14	57	22.6	316	2	rifin PRB1005w - m
15	57	22.6	725	1	A-type inclusion p
16	57	22.6	726	1	A-type inclusion p
17	57	22.6	1284	1	phosphoprotein pho
18	55	21.8	305	2	rifin PRB1015w - m
19	55	21.8	348	2	hypothetical prote
20	55	21.8	623	2	hypothetical prote
21	55	21.8	795	1	heat shock protein
22	55	21.8	822	2	hypothetical prote
23	54.5	21.6	1021	2	hypothetical prote
24	54.5	21.6	1259	2	hypothetical prote
25	54.5	21.6	1259	2	hypothetical prote
26	54.5	21.6	2715	2	eyelid - fruit fly
27	54	21.4	291	2	cyclin DI - zebra
28	54	21.4	314	2	cyclin A-type (clo
29	54	21.4	643	2	Ro autoantigen 60K

30	54	21.4	880	2	F75103	probable purine NR
31	53.5	21.2	67	2	T15592	hypothetical prote
32	53.5	21.2	301	2	T54209	hypothetical prote
33	53.5	21.2	1642	2	T08880	NMDA receptor-bind
34	53	21.0	622	2	E69006	glutamate synthase
35	53	21.0	702	2	S46854	A28L protein - var
36	53	21.0	702	2	B72167	hypothetical prote
37	53	21.0	702	2	T28570	hypothetical prote
38	53	21.0	795	2	T50255	108K heat shock pr
39	52.5	20.8	91	1	S00219	ubiquinol--cytochr
40	52.5	20.8	96	1	XLH0A	colipase A precurs
41	52.5	20.8	275	2	B81430	probable prephenat
42	52.5	20.8	425	2	T18592	hypothetical prote
43	52.5	20.8	600	2	T18593	hypothetical prote
44	52.5	20.8	1172	1	TSHP2	thrombospondin 2 p
45	52	20.6	33	2	A41822	antimicrobial pept

ALIGNMENTS

RESULT 1
S08059
alpha-globulin type B precursor (tandem 1) - upland cotton (fragment)
N:Alternate names: seed storage protein
C:Species: Gossypium hirsutum (upland cotton)
C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 30-Sep-1993
C:Accession: S08059
R:Chlan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.
Plant Mol. Biol. 9, 533-546, 1987
A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.
A:Reference number: S06398
A:Accession: S08059
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-509 <CHL>
C:Superfamily: glycylcin

Query Match 30.8% Score 77.5; DB 2; Length 509;
Best Local Similarity 41.2% Pred. No. 0.11;
Matches 14; Conservative 9; Mismatches 10; Indels 1; Gaps 1;

OY 5 ROEYECCKRQCQMLETSGOMRCVSOCDKRFED 38
Db 3 QRRTECCQECRQOE-ERQOPCQCRCKRFEQ 35

RESULT 2
FMCNAB
alpha-globulin B precursor (clone C72) - upland cotton
N:Alternate names: seed storage protein; vicilin precursor
C:Species: Gossypium hirsutum (upland cotton)
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C:Accession: A30838; S06911
R:Chlan, C.A.; Pyle, J.B.; Legocki, A.B.; Dure III, L.
Plant Mol. Biol. 7, 475-489, 1986
A:Title: Developmental biochemistry of cottonseed embryogenesis and germination XVIII
A:Reference number: A30838
A:Accession: A30838
A:Molecule type: mRNA
A:Residues: 1-588 <CHL>
A:Cross-references: GB:M16891; NID:q167374; PIDN:AAA3071.1; PID:q167375
A:Experimental source: var. Coker 201
R:Chlan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.
Plant Mol. Biol. 9, 533-546, 1987
A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.
A:Reference number: S06398
A:Accession: S06911
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-81 <CH2>
C:Comment: This is a seed storage protein.

C:Superfamily: glycinin
C:Keywords: glycoprotein; seed; storage protein
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-588/Product: alpha-globulin storage protein #status predicted <MAT>
F:411/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	30.8%	Score 77.5;	DB 1;	length 588;
Best Local Similarity	40.58;	Pred. No. 0.13;		
Matches 15;	Conservative 9;	Mismatches 12;	Indels 1;	Gaps 1

```

Oy      2  EFDROEIEECKRQCMQLETSGOMRRCVSQCCKRFEED 38
          | :: |||::: | | | : | : |||:::
Db      81  EDPQRREIECCQCEERQGE-ERQRPQCQQRCLKRFEGE 116

```

```

RESULT      3
S06398
alpha-globulin type A precursor - upland cotton
N:Alternate names: seed storage protein
C:Species: Gossypium hirsutum (upland cotton)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 30-Sep-1993
C:Accession: S06398
R:Chan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.
Plant Mol. Bioi. 9, 533-546, 1987
A>Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX. Se
A:Reference number: S06398
A:Accession: S06398
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-605 <CHL>
C:Superfamily: glycinin
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-605/Product: alpha-globulin type A #status predicted <MAT>

```

Query Match	28.68;	Score 72;	DB 2;	Length 605;
Best Local Similarity	42.58;	Pred. NO. 0.57;		
Matches 17;	Conservative 9;	Mismatches 10;	Indels 4;	Gaps 3

```

Oy      1 SEDP-QDEEBCRRCMOLETSGOMRCVSOCDKFEEDI 39
        | | : | : : | | | | | : : | | :
Db      33 SEDPQQRREDCKRKC-QLETRGQYEQ-DKCEDRSTQL 69

```

RESULT 4
S22477
vicillin precursor - cacao
C.Species: Theobroma cacao (cacao)
C.Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999
C.Accession: S22477; S22478; S18105; S22050
R:McHenry, L.; Filtz, P. J.
Plant Mol. Biol. 18, 1173-1176, 1992
A.Title: Comparison of the structure and nucleotide sequence of vicillin genes of cocoa
A.Reference number: S22477; MUID:92288309
A.Accession: S22477
A.Molecule type: DNA
A.Residues: 1-566 <MC>
A:Cross-references: EMBL:X62625
A.Accession: S22478
A.Molecule type: mRNA
A.Residues: 1-452 <MC2>
A:Cross-references: EMBL:X62626
C.Genetics:
A:Introns: 211/1; 269/3; 296/3; 391/3; 502/1
C:Superfamily: glycinin
C:keywords: seed, storage protein
E:1-24/domain: signal sequence #status predicted <Sig>
E:25-566/Product: vicillin #status predicted <Mat>

Query Match	28.0%;	Score 70.5;	DB 2;	Length 566;
Best Local Similarity	32.4%;	Pred. No. 0.8;		

	Matches	11; Conservative	11; Mismatches	11; Indels	1; Gaps
QY	5	ROVEYECKRQCMQLET-SGQMRRCVSQCDKREE	37		
	11; 11; 11; 11; 11	1	: : 1	: : : 1	
Db	39	ROVEYEQCQRCESEATEEREYEQCQRCREKE	72		

```

RESULT      5
JC5557
  arginine/glutamate-rich 6.5K polypeptide - smooch loofah
C:Species: Luffa cylindrica (smooch loofah)
C:Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 07-May-1999
C:Accession: JC5557
R:Kimura, M.; Park, S.S.; Sakai, R.; Yamasaki, N.; Funatsu, G.
  Biosci. Biotechnol. Biochem. 61, 984-988, 1997
A:Title: Primary structure of 6.5K-arginine/glutamate-rich polypeptide from the seeds
A:Reference number: JC5557; MUID:97357433
A:Accession: JC5557
A:Molecule type: protein
A:Residues: 1-47 <KIM>
A:Experimental source: seed
C:Comment: This protein is a storage protein which provides nitrogen and carbon reser
P:12-33,16-29/Disulfide bonds: #status predicted

```

Query Match	26.8%;	Score 67.5;	DB 2;	Length 47;
Best Local Similarity	41.2%;	Pred. No. 0.19;		
Matches 14;	Conservative 4;	Mismatches 15;	Indels 1;	Gaps 1;

```
OY      5 RQEFEECKRQCMOLETS-GQMRRCVSQCDKREE 37
          | | | : : | | | : | | |
DB      6 RTEACRACVRQVAEHGVERQRRRCQVCEKRLRE 39
```

```

RESULT      6
T44430
protein PV100 [imported] - winter squash
C:Species: Cucurbita maxima (winter squash)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T44430
R:Yamada, K.; Shimada, T.; Kondo, M.; Nishimura, M.; Hara-Nishimura, I.
J. Biol. Chem. 274, 2563-2570, 1999
A:Title: Multiple functional proteins are produced by cleaving Asn-Gln bonds of a single
A:Reference number: F22767; MUID:99107919
A:Accession: T44430
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-810 <YAM>
A:Cross-references: EMBL:AB019195; NID:G3808061; PIDN:BAA34056.1; PID:G3808062

```

Query Match	26.6%;	Score 67;	DB 2;	Length 810;
Best Local Similarity	34.1%;	Pred. No. 2.8;		
Matches	15;	Conservative	6;	Mismatches 15;
				Indels 8;
				Gaps 2;

```
QY 5 RQYEYECKRQCMQLETS-GQMRKCVSQCCKRFE-----EDID 40
      | ||| : : | | : | : | : | : | : | : | : | : |
Db 75 RALEYEVCLRCQVAERGVGGQQRKCEQYVCEERLREREGRGEDVD 118
```

```

RESULT      7
TSHUP1
Thrombospondin 1 precursor - human
C/Species: Homo sapiens (man)
C/Date: 23-Aug-1987 #sequence,revision 03-Aug-1995 #text,change 13-Aug-1999
C/Accession: A26155; A34274; A30140; A25812; A05172; A42927
R/Lawler, J.; Hynes, R.O.
J. Cell Biol. 103, 1635-1648, 1986
A>Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple
A/Reference number: A26155; MUID:87057617
A/Accession: A26155
A/Molecule type: mRNA
A/Residues: 1-1170 <LAW>

```



```

N:Alternate names: protein F2K15.210
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T45840
R:Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Lemcke,
submitted to the Protein Sequence Database, January 2000
A:Reference number: 223015
A:Accession: T45840
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-554 <RIB>
A:Cross-references: EMBL:AL132956
A:Experimental source: cultivar Columbia; BAC clone F2K15
C:Genetics:
A:Map position: 3
A:Introns: 53/2; 98/3; 115/2; 133/2
A:Note: F2K15.210

Query Match          25.4%; Score 64; DB 2; Length 554;
Best Local Similarity 28.2%; Pred. No. 4.5;
Matches 11; Conservative 10; Mismatches 18; Indels 0; Gaps 0;

Oy      5 ROEYEECKKRCQMOLETSGOMRCVSCDCKRFEEDIDWSK 43
      1:111 :111 :1: : : :1:
Db      148 RKEVRLRQCKRLQKHNNGTPLKLYNGSETTIDEYDWER 186

RESULT 10
JO1730
62K sucrose-binding protein precursor - soybean
C:Species: Glycine max (soybean)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 26-Aug-1999
C:Accession: JO1730
R:Grimes, H.D.; Overvoorde, P.J.; Ripp, K.; Franceschi, V.R.; Hiltz, W.D.
Plant Cell 4, 1561-1574, 1992
A:Title: A 62-kD sucrose binding protein is expressed and localized in tissues actively
A:Reference number: JO1730; MUID:93104680
A:Accession: JO1730
A:Molecule type: mRNA
A:Residues: 1-524 <GRI>
A:Cross-references: GB:106038; NID:g1431744; PIDN:AAB03894.1; PID:g170064
C:Superfamily: glycolin
C:Keywords: sugar transport
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-524/Product: 62K sucrose-binding protein #status predicted <MAT>

Query Match          25.0%; Score 63; DB 2; Length 524;
Best Local Similarity 38.2%; Pred. No. 5.6;
Matches 13; Conservative 4; Mismatches 15; Indels 2; Gaps 1;

Oy      2 EFDROEYEECKKRCQMOLE--TSGOMRCVCSODCK 33
      1: :111111:111111:
      34 EEDPELVTCRKHQCQQQQQYTEGDKRVCLOQSDCR 67

RESULT 11
T28432
variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)
N:Alternate names: erythrocyte membrane binding protein 1 (EMP1)
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T28432
R:Su, X.Z.; Heattole, V.M.; Wertheimer, S.P.; Guinet, F.; Herrfeldt, J.A.; Peterson, D.S.
Cell 82, 89-100, 1995
A:Title: The large diverse gene family var encodes proteins involved in cytoadherence an
A:Reference number: Z20487; MUID:95330813
A:Accession: T28432
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-3078 <SUX>
A:Cross-references: EMBL:L40608; NID:g9886374; PID:g9886375; PIDN:AAA75396.1

```

[illegible]

